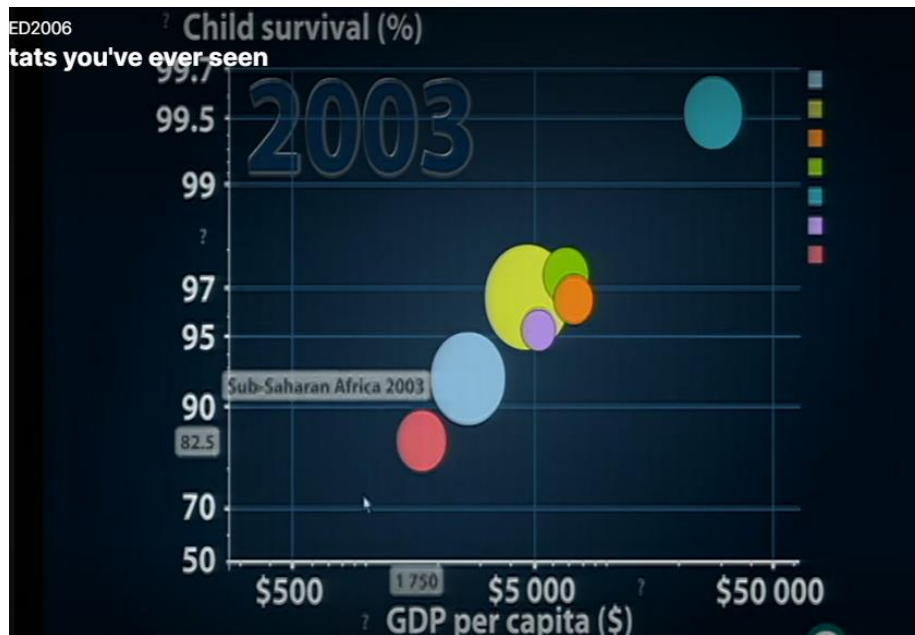


Rish of Aggregation and Generalization

Author: Eva Wang
Written on: Sept 24, 2020

Today I watched the presentation by Hans Rosling about mortality ([Hans Rosling presentation on "The Best Stats You've Ever Seen"](#)). It is one of the best presentations about visualization and statistics. In below, I will explain my two takeaways.

My biggest takeaway is the **danger of aggregation**. If we aggregate the income and mortality by the region, it seems that the richer the people are, the higher chance for the children to survive. The y axis represents the mortality rate. Africa is on the bottom left side.



A snapshot of Han's presentation -1

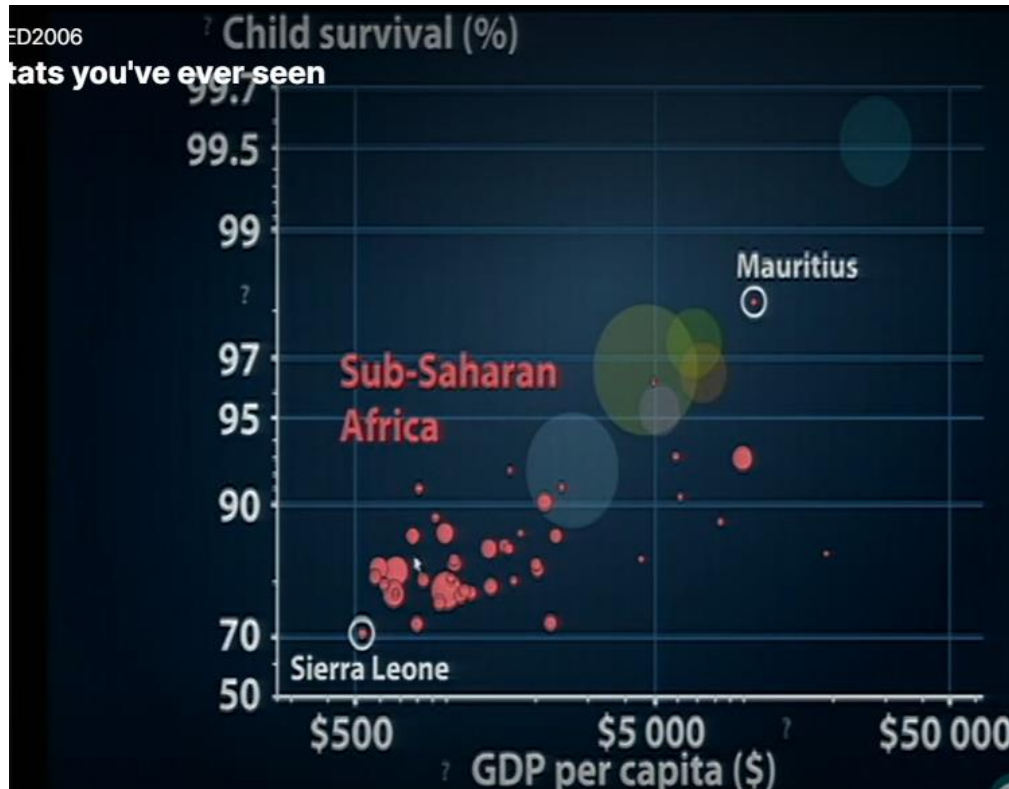
However, aggregation could deceive us. After drilling down the data, Has found that one region (represented by the purple line) shows a negative relationship between GDP per capita and Child Survival rate.



A snapshot of Han's presentation -2

This will remind me to be cautious when analyzing the correlational relationship between two factors using the average numbers.

Another takeaway is the **risk of generalization**. When drilling down/smashing the bubble representing Africa, I saw that some African countries are not poor. On the upper left side of the below snapshot, I saw a red bubble with more than \$ 5000 GDP per capita.



A snapshot of Han's presentation -3

I recalled the lectures I often heard when I was little: "You should not waste your food. African kids do not even have food to eat." Now I should reply: Africa does have quite a few developing countries, but not all African countries are poor.